(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Burcau

(43) International Publication Date
1 August 2002 (01.08.2002)

PCT

(10) International Publication Number
WO 02/059331 A1

(51) International Patent Classification[7]: C12N 15/80 // 5/10, 15/01

(21) International Application Number: PCT/DK02/00050

(22) International Filing Date: 24 January 2002 (24.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2001 00123 24 January 2001 (24.01.2001) DK
60/265,300 31 January 2001 (31.01.2001) US

(71) Applicant *(for all designated States except US)*: NOVOZYMES A/S [DK/DK]; Krogshøjvej 36, DK-2880 Bagsværd (DK).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: VIND, Jesper [DK/DK]; Hejrebakken 20, DK-3550 Værløse (DK).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IN VIVO MUTATION AND RECOMBINATION IN FILAMENTOUS FUNGI

(57) Abstract: A filamentous fungal cell having a mutator phenotype or having a reduced nucleotide mismatch repair efficiency as compared to a wildtype cell, a vector and method for making such a cell, and a method for making diverse polynucleotide libraries in a mutator phenotype cell.

WO 02/059331 A1

In viv mutati n and recombinati n in filament us fungi

FIELD OF INVENTION

A filamentous fungal cell having a mutator phenotype or having a reduced nucleotide mismatch repair efficiency as compared to a wildtype cell, a vector and method for making such a cell, and a method for making diverse polynucleotide libraries in a mutator phenotype cell.

BACKGROUND OF THE INVENTION

The mismatch repair system is a system within cells which recognises mismatches in newly synthesised duplex polynucleotides.

The mismatch repair system may correct nucleotide mismatches e.g. by using the methylated "old" strand as template, or alternatively the system may mediate degradation of the double-stranded DNA sequences that comprise the mismatches.

Independently of the precise mechanism, the end result will be that the mismatch repair system will limit the genetic diversity within a cell, diversity here being represented as duplex DNA sequences which comprise mismatches.

A duplex DNA sequence comprising a single mismatch represents a diversity of two, since there are two different DNA sequences within the cell. If the mismatch repair system corrects the single mismatch the result is a diversity of only one.

Likewise, if the mismatch repair system mediates the degradation of such a duplex DNA sequence, the diversity will be lost. Consequently, if duplex DNA sequences comprising mismatches represent a DNA library of interest, then the diversity of this library may be limited when transformed (introduced) into cells having an active mismatch repair system.

A function of the mismatch repair system in a living cell is to maintain genetic stability by minimizing the number of mutations arising e.g. from errors in DNA replication. In human diseases such as colorectal cancer, it has been found that the cause of the disease can be mutations in the mismatch repair protein encoding msh2 genes (Alas M., Bruin R., Eyck L.T., Los G., Howell S. (1998) Faseb journal vol. 12 p. 653-663). The cancer may occur when a msh2 mutant is expressed in cells, even if there is a functional wildtype msh2 gene present as well.

Some variants of the human MSH2 protein (G674A, K675A and S676A) were made and the mutant proteins was found to bind very poorly to DNA containing mismatches and have poor ATPase activity compared to wildtype MSH protein (Whitehouse A., Parmar R., Deeble J., Taylor G., Phillips S., Meredith D., Markham A. (1996) Biochem. Biophys. Res. comm. Vol 229 p 147-153).

Dominant negative msh mutations have been identified in yeast, including the polypeptide positions G693D, K694R and S695P (Studamire B., Price G., Sugawara N., Haber J., Alani E. (1999) Mol. Cell. Biol. Vol 19 p. 7558-7567). These MSH protein mutations were shown to give an approx. 50-fold increase in overall cellular mutation frequency (measured as canavanine resistance) when compared to cells having wildtype msh genotype. The mutations G693D and K694R also reduced double stranded break repair (DSB repair) demonstrating their importance in genome recombination.

The MSH2 protein binds to the MSH6 protein forming a MSH-complex, which in turn binds to mismatch containing DNA heteroduplexes, the MSH-complex then becomes active in repairing the DNA, utilizing ATP in the process.

A dominant negative mutant MSH2 protein may bind to the MSH6 protein creating an inactive MSH-complex, which does not bind to heteroduplex DNA. The binding of the mutant MSH2 protein to MSH6 protein thus leaves less free MSH6 protein to form active MSH-complex with wildtype MSH2 protein. Another dominant negative MSH2 protein may likewise create an inactive MSH-complex, which *does* bind to the heteroduplex DNA but remains inactive e.g. due to reduced ATPase activity.

The absence of active MSH-complex in a cell leads to a reduced mismatch repair efficiency which is often referred to as a "mutator phenotype", since the observed phenotype seems to be an increase in the overall mutation frequency.

The art describes several types of cells wherein the mismatch system is inactive. WO 97/37011 describes yeast cells wherein the mismatch repair system is inactivated. WO 97/05268 describes mice cells wherein the mismatch repair system is inactivated. We have previously reported filamentous fungal cells wherein the mismatch repair system encoded by the gene *msh2* is inactive WO 00/50567.

SUMMARY OF THE INVENTION:

The problem to be solved by the present invention is to provide means and method suitable for the efficient making of diverse polynucleotide libraries in filamentous fungal cells.

A filamentous fungal cell population comprising a polynucleotide library may be used to select a polynucleotide encoding a polypeptide of interest. Polynucleotides with particular properties may also be selected, such as promoters, terminators and other regulatory elements with changed/improved properties.

The solution is based on the present inventors demonstrating herein that expression of a gene encoding a dominant negative mutant of a mismatch repair protein in a wildtype filamentous fungal cell reduces the mismatch repair repair efficiency in said cell to "mutator phenotype" level.

The mismatch repair gene that was mutated and expressed herein was previously cloned and identified by the present inventors (WO 00/50567) as encoding the protein MSH2, this was the first gene cloned which is involved in the mismatch repair system of a filamentous fungal cell.

By expressing dominant negative MSH2-mutants in filamentous fungal cells it is possible to obtain cells that are impeded in their MSH2-dependent DNA mismatch repair system and such fungal cells are highly useful for preparing diverse polynucleotide libraries through *in vivo* mutation and/or recombination of one or more polynucleotides of interest.

Accordingly, in a first aspect the present invention relates to a filamentous fungal cell comprising at least one copy of a gene encoding a dominant negative mutant of a mismatch repair protein, wherein the cell has a reduced mismatch repair efficiency as compared to a wildtype cell.

The term "dominant negative mutant" of a protein in the present context means a mutant protein which, when expressed in a cell wherein a wildtype of the same protein is also expressed, dominates the wildtype protein and effectively renders the cell having a mutant phenotype. Thus expression of a dominant negative msh2 mutant in a wildtype cell will result in a cell having a MSH2 mutant phenotype. Non-limiting examples of dominant negative MSH2-mutants are given in the examples herein (G648A, K649A, and S650A). Other dominant negative repair protein mutants are known in the art see: (Studamire et al., *vide supra*), (Whitehouse et al., *vide supra*), (Gupta and Kolodner, Novel dominant mutations in *Saccharomyces cerevisiae MSH6*, 2000, Nature Genetics 24:53-56), (Bowers et al., A mutation in the MSH6 subunit of the *Saccharomyces cerevisiae* MSH2-MSH6 complex disrupts mismatch recognition, 1999, J Biol Chem 274(23):16115-16125), and (Drotschmann et al., Mutator phenotypes of yeast strains heterozygous for mutations in the *MSH2* gene, 1999, PNAS 96:2970-2975) all of which are incorporated herein by reference.

The term "mismatch repair protein" herein means a protein involved in nucleotide mismatch repair of heteroduplex polynucleotides such as DNA. Examples of mismatch repair proteins and genes encoding such proteins are given below and include MutS, MutL, MutH, MSH2, MSH3, MSH6, MLH1, MLH3, PMS1, MutM, MutY, MutT, MutH, HexA, HexB, and GTBP/p160.

In the present invention it is anticipated that it may be an advantage to have more than one copy of the gene encoding the dominant negative mutant repair protein present in a cell, thus two, three, or four copies may provide a higher overall mutation frequency in the cell or reduce the mismatch repair efficiency to even lower levels.

The term "reduced mismatch repair efficiency" is used synonymously herein with the terms "mutator phenotype" and "increased mutation rate/frequency". A reduced mismatch

repair efficiency in a cell results in a build-up within the cell of mismatch mutations that constantly occur in all cells e.g. during replication. Even though the actual overall mutation frequency may be the same in a wildtype cell and in a mutated cell having a reduced mismatch repair efficiency, for all practical purposes in the present context, the observed phenotype of the mutant cell will appear to be that of an increased frequency of mutation. As is shown in the examples below, the frequency of mutation can be determined by performing an assay if necessary.

Several strategies are known in the art that with the present invention can be used for constructing a filamentous fungal cell of the first aspect. Genomic integration of the gene encoding the dominant negative mutant is one way. Another way is to introduce into the filamentous fungal cell a stably maintained extrachromosomal genetic element, such as a plasmid, which comprises at least one copy of the gene encoding the dominant negative mutant. More copies may provide a further advantage.

The present invention also provides an extrachromosomal genetic element suitable for creating a cell of the first aspect, wherein said element is capable of autonomous maintenance in a filamentous fungal cell, and comprises at least one copy of the gene encoding the dominant negative mutant as well as at least one copy of a polynucleotide encoding a polypeptide of interest.

Accordingly, in a second aspect the present invention relates to a recombinant vector capable of autonomous maintenance in a filamentous fungal cell, the vector comprising at least one copy of a gene encoding a dominant negative mutant of a mismatch repair protein.

The term "autonomous maintenance" is well-recognized in the art and means that the vector is properly replicated and segregated in the progeny cells during cell division, so that the vector is not lost from the cell during growth.

The present inventors anticipate that it may be advantageous to have two, three, four or more copies of the gene encoding a dominant negative mutant of a mismatch repair protein comprised in the vector of the second aspect.

Allowing a filamentous fungal cell with reduced mismatch repair efficiency to replicate polynucleotides of interest at least once provides a final polynucleotide library in the cell wherein nucleotide mismatches generated during replication have accumulated. This approach provides a library with a higher diversity as compared to a library made in a wildtype filamentous fungal cell.

"Replication" herein has the recognized meaning known in the art, wherein the two strands of a polynucleotide duplex are replicated such that two separate sets of double-stranded polynucleotides are generated, each being based on a separate strand of the two original strands. The diversity of the library may be increased by doing several cycles of

replication e.g. repeating step (b) of the third aspect of the invention, wherein the diversity increases with each replication cycle.

Introducing polynucleotides into a filamentous cell may be done according to any of the many suitable techniques, such as transformation techniques. See the general fungal textbook "Fungal Genetic" (1996, ISBN 0-8247-9544-X) for a further description of such standard techniques.

A practical example may be that single stranded oligonucleotide sequences partially homologous to chromosomal DNA sequence are placed within the cell. See Calissano et al. (Fungal genetic newsletter 43:15-16 (1995) for further description of this.

The polynucleotide of interest may be comprised in a plasmid wherein the plasmid is characterised by that it comprises a suitable replication initiating sequence and optionally a suitable selectable marker as described above. Preferably the suitable replication initiating sequence is AMA1 (Gems, D., et al. (1991, Gene 98:61-67).

Cultivating a filamentous fungal cell may be done in any of the numerous suitable known media for growing filamentous fungal cells. It is within the skilled persons general knowledge to choose such a suitable media.

Since inactivation of the mismatch repair system normally will cause accumulation of mutations on the chromosomal DNA within the cell and thereby maybe make lethal mutations to the cell the actual preferred number of replication cycles as described above will depend on how fast such potential lethal mutations arise. It is within the skilled persons general knowledge to determine how many of duplication cycles it preferred. Due to these potential lethal mutations it may be preferred that the mismatch repair system has been inactivated transitorily. Then, after suitable cycles of replication according to step (b) of the third aspect, the transitorily inactivated mismatch repair system may then be re-activated in order to avoid lethal mutations in the filamentous fungal cell during later cultivation.

Accordingly, in a third aspect the present invention relates to a method for preparing a filamentous fungal cell comprising a diverse polynucleotide library of interest, the process comprising:

a) introducing a polynucleotide of interest or a polynucleotide library of interest into a filamentous fungal cell as defined in the first aspect; and

b) cultivating the cell of (a) under conditions allowing the polynucleotide of interest or the polynucleotide library of interest to be replicated at least once.

As stated above a filamentous fungal cell of the first aspect of the invention is very suitable for making a diverse polynucleotide library of interest.

Accordingly, in a fourth aspect the present invention relates to the use of a filamentous fungal cell as defined in the first aspect for the generation of a diverse polynucleotide library of interest.

As already mentioned above, industrial production of polypeptides of interest is of considerable commercial interest. Industrially interesting polypeptides include pharmaceutical proteins and peptides, enzymes, antibodies, peptide-hormones, antimicrobial peptides, fungicidal peptides, and herbicidal peptides.

Consequently, in a final aspect the present invention relates to a process for producing a polypeptide of interest, the method comprising:

a) introducing a polynucleotide of interest or a polynucleotide library of interest into a filamentous fungal cell as defined in the first aspect;

b) cultivating the cell of step (a) under conditions allowing the polynucleotide of interest or the polynucleotide library of interest to be replicated at least once;

c) selecting a polynucleotide encoding a polypeptide of interest from the at least once replicated polynucleotides of step (b); and

d) expressing the selected polynucleotide in a suitable expression system to produce the polypeptide of interest.

The desired polynucleotide of interest may be a polynucleotide encoding any polypeptide comprising a desired technical feature, such as improved stability (shelf-life); improved temperature stability; changed/improved specific activity; pH optimum; improved wash performance in a detergent etc.

The fungal cell is cultivated in a suitable medium and under suitable conditions for screening or selecting for a cell harbouring the polynucleotide of interest having or encoding a desired characteristic. The cultivation may be performed in accordance with methods well-known in the art for screening of polynucleotide variant libraries.

The selected polynucleotide may be expressed by a suitable expression system of the art such as described elsewhere herein.

An advantage of the method of the final aspect may be that the polynucleotide of interest is selected from a filamentous fungal cell expressing the polynucleotide. Consequently, it is directly known that the polynucleotide can be expressed from a filamentous fungal cell, which is useful if it is subsequently intended to express the polypeptide of interest in large scale in a filamentous fungal cell. This may be of particular interest when the polynucleotide library encodes polypeptides of interest which are derived from filamentous fungal cells, since it is known that filamentous fungal polypeptides preferably are produced in industrial relevant high yields in filamentous fungal cells.

This is contrary to a similar selection process using e.g. a yeast cell. Here the only thing known is that the selected polypeptide is capable of being expressed in yeast and later expression a filamentous fungal cell might give problems, especially if high yields are required.

Many polypeptides are of industrial interest and are produced industrially in large scale plants, it is of industrial interest to diversify polynucleotides that encode polypeptides of interest in order to obtain libraries for later screening for polypeptides with changed and/or improved properties. As mentioned above, a cell of the first aspect of the invention is useful for making such polynucleotide libraries.

DEFINITIONS:

The following section provides further definitions of technical features in the aspects of the invention.

The term "a gene" denotes herein a gene (a DNA sequence) which encodes a polypeptide according to the art. A gene is defined as an open reading frame (ORF) starting from a start codon (normally "ATG", "GTG", or "TTG") and ending at a stop codon (normally "TAA", TAG" or "TGA").

In order to express said gene there must be elements, as known in the art, in connection with the gene, necessary for expression of the gene within the cell. Such standard elements may include a promoter, a ribosomal binding site, a termination sequence, and may be others elements as known in the art.

Suitable assays to test whether or not a filamentous fungal cell as described herein has a reduced mismatch repair efficiency (also termed a "mutator phenotype") are described in the examples below and include a Chlorate-resistance assay and a "gel shift assay" (also described in WO 97/05268, page 16,17 and 25).

The sequence identity in relation to the terms
"a DNA sequence encoding a polypeptide sequence which is at least *70% identical* to the polypeptide sequence" is determined as the degree of identity between two sequences indicating a derivation of the first sequence from the second. The identity may suitably be determined by means of computer programs known in the art such as GAP provided in the GCG program package (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711) (Needleman, S.B. and Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-453). Using GAP with the following settings for polypeptide sequence comparison: GAP creation penalty of 3.0 and GAP extension penalty of 0.1, th polypeptide encoded by an analogous DNA sequence of the invention exhibits a degree of identity preferably of at least 70%, more

preferably at least 80%, more preferably at least 90%, more preferably at least 95%, and especially at least 97% with the amino acid sequence in question.

The term "diversity" in the present context represents the number of non-identical variants of one polynucleotide that exist in a cell population or in a polynucleotide library, thus the term "diverse DNA library" or "diverse polynucleotide library" denotes herein a library of at least two non-identical polynucleotides. For many practical purposes the library is much larger. Accordingly, a diverse library preferably comprises at least 1000 different polynucleotides, more preferably at least 10000 different polynucleotides, and even more preferably at least 100000 different polynucleotides.

The filamentous fungal cell

The filamentous fungal cell as described herein includes all filamentous forms of the subdivision Eumycota and Oomycota. The filamentous fungi are characterised by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative. In a preferred embodiment, the filamentous fungal cell is a cell of a species of, but is not limited to, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Scytalidium*, *Thielavia*, *Tolypocladium*, and *Trichoderma*.

Examples of filamentous fungal cells of use in the present invention include an *Aspergillus* cell, an *Acremonium* cell, a *Fusarium* cell, a *Humicola* cell, a *Mucor* cell, a *Myceliophthora* cell, a *Neurospora* cell, a *Penicillium* cell, a *Thielavia* cell, a *Tolypocladium* cell, and a *Trichoderma* cell.

More specifically, the filamentous fungal cell is an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, or *Aspergillus oryzae* cell; a *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotricioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum* cell or a *Fusarium venenatum* cell (Nirenberg sp. nov; a *Humicola insolens* cell or a *Humicola lanuginosa* cell; a *Mucor miehei* cell; a *Myceliophthora thermophila* cell; a *Neurospora crassa* cell; a *Penicillium purpurogenum* cell; a *Thielavia terrestris* cell; or a *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

Replication initiating sequences

As used herein, the term "fungal replication initiating sequence" is defined as a nucleic acid sequence which is capable of supporting autonomous maintenance/replication of an extrachromosomal molecule e.g. a plasmid in a filamentous fungal cell, normally without structural rearrangement of the plasmid or integration into the host cell genome. The replication initiating sequence may be of any origin as long as it is capable of mediating replication initiating activity in a fungal cell. Preferably, the replication initiating sequence is obtained from a filamentous fungal cell, more preferably a strain of *Aspergillus*, *Fusarium* or *Alternaria*, and even more preferably, a strain of *A. nidulans*, *A. oryzae*, *A. niger*, *F. oxysporum* or *Alternaria altenata*.

A replication initiating sequence may be identified by methods well-known in the art. For instance, the sequence may be identified among genomic fragments derived from the organism in question as a sequence capable of sustaining autonomous replication in yeast, (Ballance and Turner, Gene, 36 (1985), 321-331), an indication of a capability of autonomous replication in filamentous fungal cells. The replication initiating activity in fungi of a given sequence may also be determined by transforming fungi with contemplated plasmid replicators and selecting for colonies having an irregular morphology, indicating loss of a sectorial plasmid which in turn would lead to lack of growth on selective medium when selecting for a gene found on the plasmid (Gems et al, Gene, 98 (1991) 61-67). AMA1 was isolated in this way. An alternative way to isolate a replication initiating sequence is to isolate natural occurring plasmids (e.g. as disclosed by Tsuge et al., Genetics 146 (1997) 111-120 for *Alternaria aternata*).

Examples of replication initiating sequences include, but are not limited to, the ANS1 and AMA1 sequences of *Aspergillus nidulans*, e.g., as described, respectively, by Cullen, D., et al. (1987, Nucleic Acids Res. 15:9163-9175) and Gems, D., et al. (1991, Gene 98:61-67).

The term "replication initiating activity" is used herein in its conventional meaning, i.e. to indicate that the sequence is capable of supporting autonomous replication of an extrachromosomal molecule, such as a plasmid or a DNA vector in a fungal cell.

The term "without structural rearrangement of the plasmid" is used herein to mean that no part of the plasmid is deleted or inserted into another part of the plasmid, nor is any host genomic DNA inserted into the plasmid.

Filamentous fungal selective marker

The term "selective pressure" is defined herein as culturing a filamentous fungal cell, containing a DNA vector containing a fungal selective marker gene operably linked to a polynucleotid sequence of interest, in the presence of an effective amount or the absence of

an appropriate selective agent. The effective amount of the selective agent is defined herein as an amount sufficient for allowing the selection of cells containing the selection marker from cells which do not contain the selection marker.

In a preferred embodiment, the fungal selective marker is selected from the group of genes which encodes a product capable of providing resistance to biocide or viral toxicity, resistance to heavy metal toxicity, or prototrophy to auxotrophs.

In a more preferred embodiment, the prototrophy is obtained from an enzyme selected from the group of metabolic pathways consisting of nucleotide synthesis, cofactor synthesis, amino acid synthesis, acetamide metabolism, proline metabolism, sulfate metabolism, and nitrate metabolism.

In an even more preferred embodiment, in the methods of the present invention the fungal selective marker is a gene selected from the group consisting of *argB* (ornithine carbamoyltransferase), *amdS* (acetamidase), *bar* (phosphinothricin acetyltransferase), *hemA* (5-aminolevulinate synthase), *hemB* (porphobilinogen synthase), *hygB* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *prn* (proline permease), *pyrG* (orotidine-5'-phosphate decarboxylase), *pyroA*, *riboB*, *sC* (sulfate adenyltransferase), and *trpC* (anthranilate synthase).

Polynucleotides and Nucleotide Sequences

The present invention also relates to polynucleotides having a nucleotide sequence which may code for a polypeptide of interest. The techniques used to isolate or clone a polynucleotide are known in the art and include isolation from genomic DNA, preparation from cDNA, or a combination thereof. The cloning of a polynucleotide of the present invention from such genomic DNA can be effected, *e.g.*, by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, *e.g.*, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York.

Other amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleotide sequence-based amplification (NASBA) may be used. The nucleotide sequence may be cloned from a filamentous fungal strain or another related organism and thus, for example, may be an allelic or species variant of the polypeptide encoding region of the polynucleotide.

The polynucleotide may be obtained by standard cloning procedures used in genetic engineering to relocate the polynucleotide from its natural location to a different site where it will be reproduced. The cloning procedures may involve excision and isolation of a desired fragment comprising the polynucleotide encoding the polypeptide, insertion of the fragment

into a vector molecule, and incorporation of the recombinant vector into a host cell where multiple copies or clones of the polynucleotide will be replicated. The polynucleotide may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof.

Modification of a polynucleotide of the present invention may be necessary for the synthesis of an encoded polypeptide, which comprises an amino acid sequence that has at least one substitution, deletion and/or insertion. These artificial variants may differ in some engineered way from the polypeptide isolated from its native source, *e.g.*, variants that differ in specific activity, thermostability, pH optimum, or the like.

It will be apparent to those skilled in the art that such modifications can be made outside the regions critical to the function of the molecule and still result in an active polypeptide. Amino acid residues essential to the activity of the polypeptide encoded by the nucleotide sequence of the invention, and therefore preferably subject to modification to achieve negative mutants, such as substitution, may be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (see, *e.g.*, Cunningham and Wells, 1989, *Science* 244: 1081-1085). In the latter technique, mutations are introduced at every positively charged residue in the molecule, and the resultant mutant molecules are tested for activity to identify amino acid residues that are critical to the activity of the molecule.

Sites of substrate-enzyme interaction can also be determined by analysis of the three-dimensional structure as determined by such techniques as nuclear magnetic resonance analysis, crystallography or photoaffinity labelling (see, *e.g.*, de Vos *et al.*, 1992, *Science* 255: 306-312; Smith *et al.*, 1992, *Journal of Molecular Biology* 224: 899-904; Wlodaver *et al.*, 1992, *FEBS Letters* 309: 59-64).

Moreover, a polynucleotide encoding a polypeptide of the present invention may be modified by introduction of nucleotide substitutions which do not give rise to another amino acid sequence of the polypeptide encoded by the nucleotide sequence, but which correspond to the codon usage of the host organism intended for production of the enzyme.

The introduction of a mutation into the polynucleotide to exchange one nucleotide for another nucleotide may be accomplished by site-directed mutagenesis using any of the methods known in the art. Particularly useful is the procedure, which utilizes a supercoiled, double stranded DNA vector with an insert of interest and two synthetic primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, extend during temperature cycling by means of e.g. *Pfu* DNA polymerase. On incorporation of the primers, a mutated plasmid containing staggered nicks is generated. Following temperature cycling, the product is treated with *Dpn*I which is specific for methylated and hemimethylated DNA to digest the parental DNA template and to select for

mutation-containing synthesized DNA. Other procedures known in the art may also be used. For a general description of nucleotide substitution, see, *e.g.*, Ford *et al.*, 1991, *Protein Expression and Purification* 2: 95-107.

Nucleic Acid Constructs

The present invention also relates to nucleic acid constructs or extrachromosomal genetic elements comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the expression of the polynucleotide in a suitable host cell under conditions compatible with the control sequences.

A nucleotide sequence encoding a polypeptide of the present invention may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the nucleotide sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying nucleotide sequences utilizing recombinant DNA methods are well known in the art.

The control sequence may be an appropriate promoter sequence, a nucleotide sequence which is recognized by a host cell for expression of the nucleotide sequence. The promoter sequence contains transcriptional control sequences, which mediate the expression of the polypeptide. The promoter may be any nucleotide sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease (WO 96/00787), as well as the NA2-tpi promoter (a hybrid of the promoters from the genes for *Aspergillus niger* neutral alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and mutant, truncated, and hybrid promoters thereof.

The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the nucleotide sequence encoding the polypeptide. Any terminator which is functional in the host cell of choice may be used in the present invention.

Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* alpha-glucosidase, and *Fusarium oxysporum* trypsin-like protease.

The control sequence may also be a suitable leader sequence, a nontranslated region of an mRNA which is important for translation by the host cell. The leader sequence is operably linked to the 5' terminus of the nucleotide sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention.

Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3' terminus of the nucleotide sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention.

Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Fusarium oxysporum* trypsin-like protease, and *Aspergillus niger* alpha-glucosidase.

The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the nucleotide sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the polypeptide. However, any signal peptide coding region which directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

Effective signal peptide coding regions for filamentous fungal host cells are the signal peptide coding regions obtained from the genes for *Aspergillus oryzae* TAKA amylase,

Aspergillus niger neutral amylase, *Aspergillus niger* glucoamylase, *Rhizomucor miehei* aspartic proteinase, *Humicola insolens* cellulase, and *Humicola lanuginosa* lipase.

The control sequence may also be a propeptide coding region that codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the genes for *Bacillus subtilis* alkaline protease (*aprE*), *Bacillus subtilis* neutral protease (*nprT*), *Saccharomyces cerevisiae* alpha-factor, *Rhizomucor miehei* aspartic proteinase, and *Myceliophthora thermophila* laccase (WO 95/33836).

Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polynucleotide or gene relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. In filamentous fungi, the TAKA alpha-amylase promoter, *Aspergillus niger* glucoamylase promoter, and *Aspergillus oryzae* glucoamylase promoter may be used as regulatory sequences. Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleotide sequence encoding the polypeptide would be operably linked with the regulatory sequence.

Expression Vectors

The present invention also relates to recombinant vectors and expression vectors/systems comprising the polynucleotide construct of the invention. The various polynucleotide and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide sequence encoding the polypeptide at such sites. Alternatively, the nucleotide sequence of the present invention may be expressed by inserting the nucleotide sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the xpression vector, the coding sequence is

located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (*e.g.*, a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

The vector may be an autonomously replicating vector, *i.e.*, a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome.

The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hygB* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenyltransferase), *trpC* (anthranilate synthase), as well as equivalents thereof.

Preferred for use in an *Aspergillus* cell are the *amdS* and *pyrG* genes of *Aspergillus nidulans* or *Aspergillus oryzae* and the *bar* gene of *Streptomyces hygroscopicus*.

The vectors of the present invention preferably contain an element(s) that permits stable integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

For integration into the host cell genome, the vector may rely on the nucleotide sequence encoding the polypeptide or any other element of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleotide sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleotide

sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleotides, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleotide sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question.

More than one copy of a nucleotide sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the nucleotide sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleotide sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleotide sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Filamentous fungal cell

"Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

In a preferred embodiment, the filamentous fungal cell is a cell of a species of, but not limited to, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium*, or *Trichoderma*.

In a more pref rred embodiment, the filamentous fungal cell is an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* or

Aspergillus oryzae cell. In another most preferred embodiment, the filamentous fungal host cell is a *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, or *Fusarium venenatum* cell. In an even most preferred embodiment, the filamentous fungal cell is a *Fusarium venenatum* (Nirenberg sp. nov.) cell. In another most preferred embodiment, the filamentous fungal cell is a *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023 and Yelton *et al.*, 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470-1474. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, *Gene* 78: 147-156 and WO 96/00787.

DETAILED DESCRIPTION OF THE INVENTION

As described in detail above, the first aspect of the invention relates to a filamentous fungal cell comprising at least one copy of a gene encoding a dominant negative mutant of a mismatch repair protein, wherein the cell has a reduced mismatch repair efficiency as compared to a wildtype cell.

A wide range of mismatch repair proteins are known in the art, as referenced elsewhere herein, the mismatch repair systems are well conserved throughout bacteria, fungi and even mammals. It is anticipated that dominant negative mutants of repair proteins homolgous to those of filamentous fungi will work in the present invention. Described in the art are such repair proteins as MutS, MutL, MutH, MSH2, MSH3, MSH6, MLH1, MLH3, PMS1, MutM, MutY, MutT, MutH, HexA, HexB, and GTBP/p160.

Accordingly, in a preferred embodiment the invention relates to a cell of the first aspect, wherein the gene encodes a dominant negative mutant of a homologue of a mismatch repair protein selected from the group consisting of MutS, MutL, MutH, MSH2, MSH3, MSH6, MLH1, MLH3, PMS1, MutM, MutY, MutT, MutH, HexA, HexB, and GTBP/p160.

It is of particular interest in the present invention to use a dominant negative mutant of a mismatch repair protein, which is endogenous to the cell, since an endogenous mutant

protein is likely to interact with the mismatch repair system of the cell in a dominant negative fashion.

A preferred embodiment of the invention relates to a cell of the first aspect, wherein the gene encodes a dominant negative mutant of a mismatch repair protein, which is endogenous to the cell.

Another preferred embodiment of the invention relates to a cell of the first aspect, wherein the gene encodes a dominant negative mutant of the mismatch repair protein MSH2, MSH3, or MSH6; preferably the gene encodes a dominant negative mutant of the mismatch repair protein MSH2.

Studamire et al. report that an important target for mutagenesis in mismatch repair proteins is the phospate binding loop consensus sequence since mutations in this loop reduces ATPase activity and results in a dominant negative mutator phenotype (Studamire B., Price G., Sugawara N., Haber J., Alani E. (1999) Mol. Cell. Biol. Vol 19 p. 7558-7567).

Accordingly a preferred embodiment of the invention relates to a cell of the first aspect, wherein the gene encodes a dominant negative mutant mismatch repair protein which has a reduced ATPase activity when compared to the wildtype protein; preferably the gene encodes a dominant negative mutant mismatch repair protein which has a more than 10% reduced, preferably 20%, or 30%, even 40%, more preferably 50%, or 60%, or even 70%, and most preferably 80% reduced ATPase activity when compared to the wildtype protein.

Another preferred embodiment of the invention relates to a cell of the first aspect, wherein the gene encodes a dominant negative mutant mismatch repair protein which is mutated in the phosphate binding loop consensus sequence.

Studamire et al. also (*vide supra*) report that the highly conserved Walker type A nucleotide binding motif in repair proteins was first identified as a target for mutagenesis.

Consequently a preferred embodiment of the invention relates to a cell of the first aspect, wherein the gene encodes a dominant negative mutant mismatch repair protein which is mutated in the Walker type A nucleotide binding motif.

Another preferred embodiment of the invention relates to a cell of the first aspect, wherein the gene encodes a dominant negative mutant mismatch repair protein which is mutated in the C-terminal half of the protein.

A way to make a filamentous fungal cell wherein it is possible to transitorily inactivate the mismatch repair system, is by expressing the gene encoding the dominant negative mutant repair protein in a transitory manner from an artificially regulated promoter, preferably an inducible and/or repressible promoter.

Accordingly, in a pr ferred embodiment the invention relates to a cell of the first aspect, wherein the gene is expressed by an inducible and/or repressible promoter.

The cell of the first aspect may be constructed by integrating the at least one copy of the gene encoding the dominant negative mutant repair protein in the genome of the cell by using standard techniques for genomic integration e.g. homologous recombination with or without antibiotic selection.

A preferred embodiment of the invention relates to a cell of the first aspect, wherein the at least one copy of the gene is integrated in the genome of the cell.

However it may be more advantageous that the at least one copy of the gene is comprised in an extrachromosomal genetic element, since this would allow the use of a wider range of filamentous fungal cells without having to go through the trouble of integrating the gene into the genome of each cell, this would of course require that the genetic element is capable of autonomous maintenance in a filamentous fungal cell.

A preferred embodiment of the invention relates to a cell of the first aspect, wherein the at least one copy of the gene is comprised in an extrachromosomal genetic element which is capable of autonomous maintenance in a filamentous fungal cell.

A well-characterized stabilizing element in filamentous fungi is the AMA1-sequence, also described in WO 00/24883.

A preferred embodiment of the invention relates to a cell of the first aspect, wherein the genetic element is a vector comprising an AMA1-sequence.

One of the key features of the present invention is the achievement of a reduced mismatch repair efficiency in a fungal cell. As mentioned elsewhere herein, the observable phenotype of a cell having a reduced mismatch repair efficiency is a "mutator phenotype". In the examples below we show how the mismatch repair efficiency can be measured by way of resulting mutation frequency.

Accordingly, a preferred embodiment of the invention relates to a cell of the first aspect, wherein the cell has at least a 20% reduced mismatch repair efficiency when compared to a wildtype cell, preferably at least 30%, or at least 40%, even at least 50%, more preferably at least 60%, or at least 70%, or at least 80%, or most preferably at least 90% reduced mismatch repair efficiency when compared to a wildtype cell.

Polypeptides are of considerably commercial interest, especially polypeptides that have new and/or improved properties compared to the existing ones. Non-limiting examples of desirable properties in this respect could be related to stability, activity, specificity, allergenicity, toxicity etc. Ways of improving and/or changing the properties of a particular polypeptide are not straightforward since the molecules are complex and many of the molecular mechanisms underlying these properties are not yet fully understood. However it is generally appreciated in the art that it is advantageous to prepare libraries of diversified

polynucleotides that may be based on the polynucleotide encoding the polypeptide of interest, and to express and screen such libraries for the desired properties.

A preferred embodiment of the invention relates to a cell of the first aspect, further comprising at least one copy of a polynucleotide encoding a polypeptide of interest.

Of particular industrial interest are enzymes, so a preferred embodiment of the invention relates to a cell of the first aspect, wherein the polypeptide is an enzyme; preferably the enzyme is an amylolytic enzyme, a lipolytic enzyme, a proteolytic enzyme, a cellulytic enzyme, an oxidoreductase or a plant cell-wall degrading enzyme, more preferably it is an enzyme with an activity selected from the group consisting of aminopeptidase, amylase, amyloglucosidase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, and xylanase.

It may be an advantage to have the at least one copy of the polynucleotide stably integrated into the chromosome of the cell, so a preferred embodiment of the invention relates to a cell of the first aspect, wherein the at least one copy of the polynucleotide is integrated in the genome of the cell.

However it may be more advantageous that the at least one copy of the polynucleotide is comprised in an extrachromosomal genetic element, since this would allow the use of a wider range of filamentous fungal cells without having to go through the trouble of integrating the gene into the genome of each cell.

Accordingly, a preferred embodiment of the invention relates to a cell of the first aspect, wherein the at least one copy of the polynucleotide is comprised in an extrachromosomal genetic element which is capable of autonomous maintenance in a filamentous fungal cell; preferably the genetic element is a vector comprising an AMA1-sequence; more preferably the vector further comprises at least one copy of the gene of the first aspect encoding a dominant negative mutant repair protein.

The art describes a plasmid pAAT56 derived from the filamentous fungus *Alternaria alternata* (I. Kaneko et al., Structural analysis of the plasmid pAAT56 of the filamentous fungus *Alternaria alternata*, 1997, Gene 203:51-57). It is anticipated that it may be an advantage in the present invention to utilize the plasmid pAAT56 as a preferred genetic element in the preceding embodiments.

Accordingly, a preferred embodiment of the invention relates to a cell of the first aspect, wherein the genetic element is a plasmid pAAT56 derivative.

As mentioned above, it is of interest to construct diversified libraries of polynucleotides in a cell, and the cell of the first aspect makes this possible through accumulation of the mismatch mutations that occur naturally in the cell e.g. during replication. In the present invention we further anticipate that it may be an advantage to increase the rate of mismatch incorporation during replication e.g. by introducing an error-prone polymerase into the cell. This would likely result in a higher diversity, since the number of accumulated mismatch mutations would likely increase in the cell. The concept has previously been described in the art, however not in combination with a filamentous fungal cell of the present invention, see e.g. WO 97/25410 or US Patent No. 6165718.

Consequently, a preferred embodiment of the invention relates to a cell of the first aspect, wherein the genetic element is replicated by an error-prone polymerase; preferably the error-prone polymerase is a reverse transcriptase.

Just like with the gene encoding the dominant negative mutant and with the polynucleotide of interest it may be an advantage if a polynucleotide encoding the error-prone polymerase is comprised in the extrachromosomal genetic element(s) of the preceding embodiments.

So, a preferred embodiment of the invention relates to a cell of the first aspect, wherein the genetic element comprises a polynucleotide encoding the error-prone polymerase.

The filamentous fungi of the present invention are many, however a cell of the genus *Fusarium* or *Aspergillus* is preferred, in particular a cell of the species *Fusarium oxysporum, Fusarium venenatum, Aspergillus nidulans, Aspergillus oryzae,* or *Aspergillus niger*

A preferred embodiment of the invention relates to a cell of the first aspect, wherein the filamentous fungal cell is a *Fusarium* or *Aspergillus* cell, preferably of the species *Fusarium oxysporum, Fusarium venenatum, Aspergillus nidulans, Aspergillus oryzae,* or *Aspergillus niger.*

As described above, one way to construct a cell of first aspect of the invention could simply be to introduce into a wildtype cell a suitable extrachromosomal element or recombinant vector comprising at least one copy of a gene encoding a dominant negative mutant of a mismatch repair protein.

So naturally a second aspect of the invention relates to a recombinant vector capable of autonomous maintenance in a filamentous fungal cell, the vector comprising at least one copy of a gene encoding a dominant negative mutant of a mismatch repair protein.

Several of the preferred embodiments of the invention relate to a vector of the second aspect, and the reasoning behind these is similar to what has already been outlined above.

A preferred embodiment of the invention relates to a vector of the second aspect, which comprises an AMA1 sequence.

Another preferred embodiment of the invention relates to a vector of the second aspect, which is derived from plasmid pAAT56.

Yet another preferred embodiment of the invention relates to a vector of the second aspect, wherein the dominant negative mutant is a mutant homologue of a mismatch repair protein selected from the group consisting of MutS, MutL, MutH, MSH2, MSH3, MSH6, MLH1, MLH3, PMS1, MutM, MutY, MutT, MutH, HexA, HexB, and GTBP/p160.

Still another preferred embodiment of the invention relates to a vector of the second aspect, wherein the dominant negative mutant is a mutant of the mismatch repair protein MSH2, MSH3, or MSH6.

One preferred embodiment of the invention relates to a vector of the second aspect, wherein the dominant negative mutant is a mutant mismatch repair protein which has a reduced ATPase activity when compared to the wildtype protein; preferably the dominant negative mutant is a mutant mismatch repair protein which has a more than 10% reduced ATPase activity when compared to the wildtype protein, preferably more than 20%, or more than 30%, even more than 40%, more preferably more than 50%, or more than 60%, or even more than 70%, and most preferably more than 80% reduced ATPase activity when compared to the wildtype protein

Another preferred embodiment of the invention relates to a vector of the second aspect, wherein the dominant negative mutant is a mutant mismatch repair protein which is mutated in the phosphate binding loop consensus sequence.

One more preferred embodiment of the invention relates to a vector of the second aspect, wherein the dominant negative mutant is a mutant mismatch repair protein which is mutated in the Walker type A nucleotide binding motif.

A preferred embodiment of the invention relates to a vector of the second aspect, wherein the dominant negative mutant is a mutant mismatch repair protein which is mutated in the C-terminal half of the protein.

Another preferred embodiment of the invention relates to a vector of the second aspect, wherein the gene encoding the dominant negative mutant is expressed by an inducible and/or repressible promoter.

One other preferred embodiment of the invention relates to a vector of the second aspect, wherein the vector further comprises at least one copy of a polynucleotide encoding a polypeptide of interest; preferably an enzyme, mor preferably an amylolytic enzyme, a lipolytic enzyme, a proteolytic enzyme, a cellulytic enzyme, an oxidoreductase or a plant cell-wall degrading enzyme; and most preferably an enzyme having an activity selected from the

group consisting of aminopeptidase, amylase, amyloglucosidase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, and xylanase.

Still further a preferred embodiment of the invention relates to a vector of the second aspect, which further comprises a polynucleotide encoding an error-prone polymerase, preferably a reverse transcriptase; preferably the vector is replicated by the error-prone polymerase.

In the third aspect, the present invention relates to a method for preparing a filamentous fungal cell comprising a diverse polynucleotide library of interest.

As mentioned above, the present invention relates to many different filamentous fungal cells, however a preferred embodiment of the invention relates to a method of the third aspect, wherein the filamentous fungal cell is a *Fusarium* or *Aspergillus* cell, preferably of the species *Fusarium oxysporum, Fusarium venenatum, Aspergillus nidulans, Aspergillus oryzae,* or *Aspergillus niger.*

Methods are well known in the art for shuffling polynucleotides (WO 98/41653; US Pat No. 6159687; US Pat No. 6117637; WO 98/41623; US Pat No.5605793; WO 95/22625; WO 97/20078) and some of these methods are carried out *in vitro* and may result in the formation of heteroduplexes. It is anticipated herein that a cell of the first aspect, a vector of the second aspect, or indeed the method of the third aspect may be useful in cloning and/or propagating such heteroduplex molecules, in order to preserve the diversity generated through the shuffling procedure.

Accordingly, a preferred embodiment of the invention relates to a method of the third aspect, wherein the polynucleotide of interest or polynucleotide library of interest comprises shuffled polynucleotides.

A preferred embodiment of the invention relates to a method of the third aspect, wherein the polynucleotide of interest or polynucleotide library of interest encodes a polypeptide of interest, preferably the an enzyme, and more preferably an amylolytic enzyme, a lipolytic enzyme, a proteolytic enzyme, a cellulytic enzyme, an oxidoreductase or a plant cell-wall degrading enzyme.

Yet another preferred embodiment of the invention relates to a method of the third aspect, wherein the polynucleotide of interest or polynucleotide library of interest is comprised in a recombinant vector as defined in the second aspect herein.

As mentioned, shuffling procedures may be used to diversify a polynucleotide library, however such diversification may also be achieved through an *in vivo* recombination event(s) between non-identical polynucleotides. Where shuffled polynucleotides are introduced into a cell of the first aspect according to the method of the third aspect, there may be several non-identical polynucleotides within one cell both before and after one or more rounds of replication. T Consequently, a preferred embodiment of the invention relates to a method of the third aspect, wherein one or more *in vivo* recombination event(s) takes place in step b) between non-identical polynucleotides of interest.

A fourth aspect of the invention relates to the use of a filamentous fungal cell as defined in the first aspect herein for the generation of a diverse polynucleotide library of interest.

The final aspect relates to a process for producing a polypeptide of interest, wherein a polynucleotide encoding a polypeptide of interest is selected from a diverse library of polynucleotides produced by a method of the third aspect and present in a cell of the first aspect, and expressing the selected polynucleotide in a suitable expression system to produce the polypeptide of interest.

Methods of Production

The present invention also relates to methods for producing a polypeptide of the present invention comprising (a) cultivating a strain, which in its wild-type form is capable of producing the polypeptide; and (b) recovering the polypeptide.

In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies,

formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide as described herein.

The resulting polypeptide may be recovered by methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

The polypeptides of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation), SDS-PAGE, or extraction (see, *e.g.*, *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

EXAMPLES:

Materials

Chemicals used as buffers and substrates were commercial products of at least reagent grade.

Example 1

Construction of an expression plasmid.

Plasmids pENI1298 and pENI1299 were described in WO 00/24883. The plasmid pENI1298 was further developed in order to decrease plasmid size (thus improve transformation), improve expression (by improving promoter), have expression in *E. coli* for improved library screening (by improving same promoter), ease cloning (by introducing unique restriction sites downstream of promoter), ease cloning (by using the Gateway cloning technology).

Plasmid pENI1960 was made using the Gateway Vector™ conversion system (Lifetechnology® cat no. 11828-019) by cutting pENI1902 with BamHI, filling the DNA ends using Klenow fragment polymerase and nucleotides (thus making blunt ends) followed by ligation to reading frame A Gateway™ PCR fragment. The cloning in the correct orientation was confirmed by sequencing.

Plasmid pENI1902 was made in order to have a promoter that works in both *E.coli* and *Aspergillus*. This was done by unique site elimination using the "Chameleon double stranded site-directed mutagenesis kit" as recommended by Stratagene .

Plasmid pENI1861 was used as template and the following primers were used as selection primers ("5´P" indicates a 5` phosphorylation of the primer):

177996 (SEQ ID 1): 5`P gaatgacttggttgacgcgtcaccagtcac
135640 (SEQ ID 2): 5´P cttattagtaggttggtacttcgag
135638 (SEQ ID 3): 5´P gtccccagagtagtgtcactatgtcgaggcagttaag

The 080399J19 primer was used as mutagenic primer to introduce a –35 and –10 promoter consensus sequence (from *E.coli*) in the *Aspergillus* expression promoter. Introduction of the mutations was verified by sequencing.

080399J19 (SEQ ID 4):
5´P gtatgtcccttgacaatgcgatgtatcacatgatataattactagcaagggaagccgtgcttgg

Plasmid pENI1861 was made in order to have the state of the art *Aspergillus* promoter in the expression plasmid, as well as a number of unique restriction sites for cloning. A PCR fragment (app. 620 bp) was made using pMT2188 (see Example 5) as template and the following primers:

051199J1 (SEQ ID 5): 5' cctctagatctcgagctcggtcaccggtggcctccgcggccgctggatccccagttgtg
1298TAKA (SEQ ID 6): 5' gcaagcgcgcgcaatacatggtgttttgatcat

The fragment was cut BssHII and Bgl II, and cloned into pENI1849 which was also cut with BssHII and Bgl II. The cloning was verified by sequencing.

Plasmid pENI1849 was made in order to truncate the pyrG gene to the essential sequences for pyrG expression, in order to decrease the size of the plasmid, thus improving transformation frequency. A PCR fragment (app. 1800 bp) was made using pENI1299 as template and the following primers:

270999J8 (SEQ ID 7): 5' tctgtgaggcctatggatctcagaac
270999J9 (SEQ ID 8): 5' gatgctgcatgcacaactgcacctcag

The PCR-fragment was cut with the restriction enzymes StuI and SphI, and cloned into pENI1298, also cut with StuI and SphI; the cloning was verified by sequencing.

Example 2

Construction of the expression plasmid pMT2188.

The *Aspergillus oryzae* expression plasmid pCaHj 483 (WO 98/00529) consists of an expression cassette based on the *Aspergillus niger* neutral amylase II promoter fused to the *Aspergillus nidulans* triose phosphate isomerase non translated leader sequence (Pna2/tpi) and the *A. niger* amyloglycosidase terminater (Tamg). Also present on the plasmid is the *Aspergillus* selective marker *amdS* from *A. nidulans* enabling growth on acetamide as sole nitrogen source. These elements are cloned into the *E. coli* vector pUC19 (New England Biolabs). The ampicillin resistance marker enabling selection in *E. coli* of this plasmid was replaced with the URA3 marker of *Saccharomyces cerevisiae* that can complement a *pyrF* mutation in *E. coli*, the replacement was done in the following way:

The pUC19 origin of replication was PCR amplified from pCaHj483 with the primers:

142779 (SEQ ID 9): 5' ttgaattgaaaatagattgatttaaaacttc

142780 (SEQ ID 10): 5' ttgcatgcgtaatcatggtcatagc

Primer 142780 introduces a *Bbu*I site in the PCR fragment. The Expand PCR system (Roche Molecular Biochemicals, Basel, Switserland) was used for the amplification following the manufacturers instructions for this and the subsequent PCR amplifications.

The URA3 gene was amplified from the general *S. cerevisiae* cloning vector pYES2 (Invitrogen corporation, Carlsbad, Ca, USA) using the primers:

140288 (SEQ ID 11): 5' ttgaattcatgggtaataactgatat

142778 (SEQ ID 12): 5' aaatcaatctattttcaattcaattcatcatt

Primer 140288 introduces an *Eco*RI site in the PCR fragment. The two PCR fragments were fused by mixing them and amplifying using the primers 142780 and 140288 in the splicing by overlap method (Horton et al (1989) Gene, 77, 61-68).

The resulting fragment was digested with *Eco*RI and *Bbu*I and ligated to the largest fragment of pCaHj 483 digested with the same enzymes. The ligation mixture was used to transform the *pyrF E.coli* strain DB6507 (ATCC 35673) made competent by the method of Mandel and Higa (Mandel, M. and A. Higa (1970) J. Mol. Biol. 45, 154). Transformants were selected on solid M9 medium (Sambrook et. al (1989) Molecular cloning, a laboratory manual, 2. edition, Cold Spring Harbor Laboratory Press) supplemented with 1 g/l casaminoacids, 500 µg/l thiamine and 10 mg/l kanamycin.

A plasmid from a selected transformant was termed pCaHj527. ThePna2/tpi promoter present on pCaHj527 was subjected to site directed mutagenises by a simple PCR approach. Nucleotide 134 – 144 was altered from GTACTAAAACC to CCGTTAAATTT using the mutagenic primer 141223. The resulting plasmid was termed pMT2188.

Primer 141223 (SEQ ID 13):

5' ggatgctgttgactccggaaatttaacggtttggtcttgcatccc

Nucleotide 423 – 436 was altered from ATGCAATTTAAACT to CGGCAATTTAACGG using the mutagenic primer 141222.

Primer 141222 (SEQ ID 14):

5' ggtattgtcctgcagacggcaatttaacggcttctgcgaatcgc

Example 3

Construction of three MSH2 dominant variants.

Five different PCR reactions using p418Msh2 (WO 00/50567) as template along with different primers were run using PWO polymerase and buffer as recommended by the manufacturer (Boehringer Mannheim™) in a standard pcr reaction (94°C 5 min, 25*(94°C 20 sec., 50°C 20 sec, 72°C 90 sec.), 72°C 5 min).

Reaction 1:

Primer 19671 (SEQ ID NO.15): 5'ctcccttctctgaacaataaaccc; and

Primer 220900j4 (SEQ ID NO.16): 5'tcccatgttaggaccagtaatgat

Reaction 2 (mutation G648A):

Primer 991213j5 (SEQ ID NO.17):

5'cctctagatctcgagctcggtcaccggtggcctccgcggccgctgcgccaggtgtcagtcaccctc; and

Primer 220900j1 (SEQ ID NO.18): 5'actggtcctaacatgggagcgaaatcgacttatattcgc

Reaction 3 (mutation K649A):

Primer 991213j5 (SEQ ID NO.17); and

Primer 220900j2 (SEQ ID NO.19): 5'actggtcctaacatgggaggtgcgtcgacttatattcgc

Reaction 4 (mutation S650A):

Primer 991213j5 (SEQ ID NO.17); and

Primer 220900j3 (SEQ ID NO.20): 5'actggtcctaacatgggaggtaaagcgacttatattcgc

Reaction 11 (wt):

Primer 991213j5 (SEQ ID NO.17); and

Primer 16062000j2 (SEQ ID NO.21):

5'ggggaccactttgtacaagaaagctgggtcctagatctcgagctcggtcac

The PCR fragments were separated by electrophoresis and purified from an 1.5 % agarose gel, and four additional PCR reactions were carried out using PWO polymerase and buffer as described by the manufacturer (Boehringer Mannheim™) in a standard PCR reaction (94°C 5 min, 25*(94°C 20 sec., 50°C 20 sec, 72°C 90 sec.), 72°C 5 min) using the purified PCR fragments as templates and the following primers:

Primer 160600J2 (SEQ ID NO.21); and

Primer 19072000J1 (SEQ ID NO.22):

5'ggggacaagtttgtacaaaaaagcaggcttctctgaacaataaaccccac

Reaction 12: with reaction 1 and 2 as template.

Reaction 13: with reaction 1 and 3 as template.

Reaction 14: with reaction 1 and 4 as template.

Reaction 17: with reaction 1 and 11 as template.

The four PCR reactions above run and PCR fragments were purified after electrophoresis from a 1.0% agarose gel. The PCR fragments were cloned into pENI1960 (10070) using the Gateway™ technology (Lifetechnology®). The "BP" cloning reaction was made by mixing the PCR fragment with BP reaction buffer, pDONR201 vector and BP clonase mix. Plasmid pENI1960 was cut with ScaI (in order to cleave in the ccdB gene), and mixed with the BP reaction and "LR" clonase mix (as recommended by Lifetechnology®) and transformed into *E. coli* DH10b (Life-Technology™ cat. no.18290-015).

The following four plasmids were constructed:

pENI2039: containing the wildtype msh2 gene in pENI1960 (PCR reaction 17).

pENI2040: containing the msh2 mutant G648A in pENI1960 (PCR reaction 12).

pENI2041: containing the msh2 mutant K649A in pENI1960 (PCR reaction 13).

PENI2042: containing the msh2 mutant S650A in pENI1960 (PCR reaction 14).

The 4 plasmids pENI2039-2042 were sequenced across the entire msh2 gene in order to confirm the introduction of the mutation as well as the lack of other unintended mutations.

Example 4

Transformation of *Aspergillus oryzae*. Induction of inactive dominant MSH protein and test of *In Vivo* mutation rates in *A. oryzae* tranformants is demonstrated in the following.

The plasmids pENI2039, pENI2040, pENI2041 and pENI2042 were constructed as outlined above from the parental plasmid pENI1960 which is derived from plasmid pENI1298 (WO 00/24883). pENI1298 has the ability to replicate autonomously in *Aspergillus* and gives equivalent expression levels between independent transformants, when measuring the activity of a reporter gene.

Thus the plasmids pENI2039, pENI2040, pENI2041 and pENI2042 replicate autonomously in *Aspergillus* and give rise to production of similar levels of MSH2 protein from the corresponding encoding msh2-genes comprised in the plasmids.

Approximately 1 microgram of the following 5 plasmids: pENI2039, pENI2040, pENI2041, pENI2042 and pENI1902 were transformed into *A. oryzae* Jal250 (JaL250 is a derivative of *A. oryzae* A1560 wherein the *pyrG* gene has been inactivated, as described in WO 98/01470; transformation protocol as described in WO 00/24883).

The transformants were incubated for 4 days at 37°C. A spore-suspension was made from the Jal250 transformants and plated onto minimal plates containing 2 % maltose and urea (2 % maltose, 10 mM Urea, 2 % agar, 2 % salt solution, 24 % sorbitol) as described in (WO 98/01470) in order to induce the promoter upstream of the msh2 gene, and to be able to cultivate transformants despite mutations in the *nia*-genes of *A. oryzae* Jal250.

The transformants are expected to have a decreased level of active MSH-complex due to production of inactive dominant MSH2 protein, thus leading to an increase in *In Vivo* overall mutation frequency.

The plates were incubated for 4 days at 37°C before a new spore-suspension was made. Dilutions were made of the spore-suspensions and plated both on minimal plates (1% glucose, 0.5 mM NH_4, 2 % agar, 2% salt solution, 0.01 % triton) where all the transformants are capable of growth, and on chlorate-containing plates (5% chlorate, 1% glucose, 0.5 mM NH_4, 2 % agar, 2% salt solution, 0.01 % triton), where only transformants mutated in the *nia*-genes that are involved in the nitrate metabolism can grow.

Chlorate is toxic to an *A. oryzae* cell if metabolised by the *nia*-gene products. Hence a high frequency of chlorate-resistant cells is due to a high *In Vivo* mutation frequency.

The Chlorate resistance frequency of the transformants is shown in table 1 (number of chlorate resistant divided by number of transformants grown on minimal plates):

Transformant-plasmid	Chlorate res. freq.
pENI1902 (control)	$5 * 10^{-6}$
pENI2039 (wildtype MSH2)	$8 * 10^{-6}$
pENI2040 (G648A MSH2)	$3 * 10^{-3}$
pENI2041 (K649A MSH2)	$5 * 10^{-4}$
pENI2042 (S650A MSH2)	$2 * 10^{-3}$

Table 1. Chlorate resistance frequency of the *A. oryzae* transformants carrying the indicated plasmids (see text for details).

As is clear from the results shown in table 1, there is only an insignificant difference between the control (pENI1902) and wildtype MSH2 (pENI2039) as would be expected. Thus an increase in the MSH2 protein level as such makes no difference in *In Vivo* mutation frequency.

However a 100-fold (pENI 2041: K649A) to 500-fold (pENI2040: G648A, and pENI2042: S650A) increase in Chlorate resistance frequency is seen, when expressing the inactive dominant mutant msh2-genes present in pENI2040-2041-2042, which is due to an increased *In Vivo* mutation frequency. Thus, expressing a dominant inactive variant of MSH2 protein in filamentous fungi, such as *Aspergillus oryzae* is a desirable way to impede or inactivate the DNA repair system. It is easy and fast to handle. It does not *per se* require any modification of the host genome. It is independent of multiple genomic copies of homologous msh2 genes. It is easy to get rid of the dominant inactive variant, as opposed to deleting or mutating the genomic msh2 gene(s), because the plasmid is lost if there is no selective pressure to maintain it in the fungal host cell. It is easy to regulate the level of expression by regulation the promoter, which in present example is induced by maltose.

Example 5

A gel shift assay may also be suitable for confirming that a filamentous fungal cell as described herein has reduced mismatch repair efficiency.

The principle in this gel shift assay is that cell extracts are prepared of both (a) a filamentous fungal cell wherein the gene, as described herein, involved in the mismatch repair system is mutated; and (b) the corresponding filamentous fungal cell wherein the gene is

NOT mutated. Thes extracts are then bound to/mixed with oligonucleotides containing the base-pair mismatched G:T; G:A; G:G; A:C, and an extrahelical TG dinucleotide and run on a nondenaturing gel.

If the gel shift assay demonstrates that the control filamentous fungal cell wherein the gene is NOT inactivated comprises any protein(s) which binds to any of above mentioned oligonucleotides and these binding protein(s) is NOT seen in the filamentous fungal cell wherein the gene involved in the mismatch repair system is mutated, then it is a confirmation that the mismatch repair protein in the latter is a mutant.

The mismatch repair protein may be mutated in such a way that it binds to heteroduplex DNA just like wildtype protein, however it remains inactive due to a loss of e.g. ATPase activity. In that case the above assay may be modified by addition of ATP to the reaction after confirming that both proteins bind to the DNA. Addition of ATP to a DNA bound wildtype protein results in the release of the protein from the DNA, whereas an ATPase negative mutant remains bound. A second gel-shift assay can confirm such a result.

Experimental procedure:

Preparation of cell extracts are performed as described in Nagata et al. (Mol. Gen Genet (1993) 237:251-260; See Materials and Methods).

Annealing of oligonucleotides, binding of cell extracts to duplex oligonucleotides containing mismatched, and nondenaturing polyacrylamide gelelectrophoresis are performed essentially as described (Stephenson and Karran; Selective binding to DNA base pair mismatches by proteins from human cells; J. Biol. Chem. 264:2177-21182 (1989)).

However, gelelectrophoresis is performed in TAE buffer rather than in TBE buffer. To obtain duplex oligonucleotides, the oligonucleotide U is radiolabelled and annealed with any of the unlabelled oligonucleotides L-G.T, L-G.A, L-G.C, L-A.C, L-T.G, or L-HOM. Oligonucleotide sequences are derived from Aquilina et al. Proc. Natl. Acad. Sci. USA 91:8905-8909 (1994).

U (SEQ ID NO 23): 5´ gggaagctgccaggccccagtgtcagcctcctatgctc

L-G.T (SEQ ID NO 24): 5´ gagcataggaggctgacattggggcctggcagcttccc (resulting in a G.T mismatch)

L-G.A (SEQ ID NO 25): 5´ gagcataggaggctgacaatggggcctggcagcttccc (resulting in a G.A mismatch)

L-G.G (SEQ ID NO 26): 5´ gagcataggaggctgacagtggggcctggcagcttccc (resulting in a G.G mismatch)

L-A.C (SEQ ID NO 27): 5´ gagcataggaggctgacaccggggcctggcagcttccc (resulting in a A.C mismatch)

L-T.G (SEQ ID NO 28) :

5´ gagcataggaggctgacactgtggggcctggcagcttccc

(resulting in an extrahelical TG dinucleotide)

L-HOM (SEQ ID NO 29): 5´ gagcataggaggctgacaccggggcctggcagcttccc (resulting in a homoduplex).

In all assays, a twofold excess of unlabelled homoduplex competitor oligonucleotide is included.

Other assays for testing whether a cell comprising a dominant negative mutant mismatch repair protein and as such has reduced mismatch repair efficiency are described in the art cited herein.

L-G.G (SEQ ID NO 26): 5´ gagcataggaggctgacagtggggcctggcagcttccc (resulting in a G.G mismatch)

L-A.C (SEQ ID NO 27): 5´ gagcataggaggctgacaccggggcctggcagcttccc (resulting in a A.C mismatch)

L-T.G (SEQ ID NO 28) :

5´ gagcataggaggctgacactgtggggcctggcagcttccc

(resulting in an extrahelical TG dinucleotide)

L-HOM (SEQ ID NO 29): 5´ gagcataggaggctgacaccggggcctggcagcttccc (resulting in a homoduplex).

In all assays, a twofold excess of unlabelled homoduplex competitor oligonucleotide is included.

Other assays for testing whether a cell comprising a dominant negative mutant mismatch repair protein and as such has reduced mismatch repair efficiency are described in the art cited herein.

CLAIMS

1. A filamentous fungal cell comprising at least one copy of a gene encoding a dominant negative mutant of a mismatch repair protein, wherein the cell has a reduced mismatch repair efficiency as compared to a wildtype cell.

2. The cell of claim 1, wherein the gene encodes a dominant negative mutant of a homologue of a mismatch repair protein selected from the group consisting of MutS, MutL, MutH, MSH2, MSH3, MSH6, MLH1, MLH3, PMS1, MutM, MutY, MutT, MutH, HexA, HexB, and GTBP/p160.

3. The cell of claim 1 or 2, wherein the gene encodes a dominant negative mutant of a mismatch repair protein, which is endogenous to the cell.

4. The cell of any of claims 1-3, wherein the gene encodes a dominant negative mutant of the mismatch repair protein MSH2, MSH3, or MSH6.

5. The cell of any of claims 1-4, wherein the gene encodes a dominant negative mutant of the mismatch repair protein MSH2.

6. The cell of any of claims 1-5, wherein the gene encodes a dominant negative mutant mismatch repair protein which has a reduced ATPase activity when compared to the wildtype protein.

7. The cell of claim 6, wherein the gene encodes a dominant negative mutant mismatch repair protein which has a more than 10% reduced, preferably 20%, or 30%, even 40%, more preferably 50%, or 60%, or even 70%, and most preferably 80% reduced ATPase activity when compared to the wildtype protein.

8. The cell of any of claims 1-7, wherein the gene encodes a dominant negative mutant mismatch repair protein which is mutated in the phosphate binding loop consensus sequence.

9. The cell of any of claims 1-7, wherein the gene encodes a dominant negative mutant mismatch repair protein which is mutated in the Walker type A nucleotide binding motif.

10. The cell of any of claims 1-7, wherein the gene encodes a dominant negative mutant mismatch repair protein which is mutated in the C-terminal half of the protein.

11. The cell of any of claims 1-10, wherein the gene is expressed by an inducible and/or repressible promoter.

12. The cell of any of claims 1-11, wherein the at least one copy of the gene is integrated in the genome of the cell.

13. The cell of any of claims 1-11, wherein the at least one copy of the gene is comprised in an extrachromosomal genetic element which is capable of autonomous maintenance in a filamentous fungal cell.

14. The cell of claim 13, wherein the genetic element is a vector comprising an AMA1-sequence.

15. The cell of any of claims 1-14, wherein the cell has at least a 20% reduced mismatch repair efficiency, preferably at least 30%, or 40%, even 50%, more preferably at least 60%, or 70%, or 80%, or most preferably at least 90% reduced, when compared to a wildtype cell.

16. The cell of any of claims 1-15, further comprising at least one copy of a polynucleotide encoding a polypeptide of interest.

17. The cell of claim 16, wherein the polypeptide is an enzyme.

18. The cell of claim 17, wherein the enzyme is an amylolytic enzyme, a lipolytic enzyme, a proteolytic enzyme, a cellulytic enzyme, an oxidoreductase or a plant cell-wall degrading enzyme, preferably an enzyme with an activity selected from the group consisting of aminopeptidase, amylase, amyloglucosidase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, and xylanase.

19. The cell of any of claims 16-18, wherein the at least one copy of the polynucleotide is integrated in the genome of the cell.

20. The cell of any of claims 16-18, wherein the at least one copy of the polynucleotide is comprised in an extrachromosomal genetic element which is capable of autonomous maintenance in a filamentous fungal cell.

21. The cell of claim 20, wherein the genetic element is a vector comprising an AMA1-sequence.

22. The cell of claim 21, wherein the vector further comprises at least one copy of the gene of claim 1.

23. The cell of claim 20, wherein the genetic element is a plasmid pAAT56 derivative.

24. The cell of any of claims 20-23, wherein the genetic element is replicated by an error-prone polymerase.

25. The cell of claim 24, wherein the error-prone polymerase is a reverse transcriptase.

26. The cell of claim 24 or 25, wherein the genetic element comprises a polynucleotide encoding the error-prone polymerase.

27. The cell of any of claims 1-26, wherein the filamentous fungal cell is a *Fusarium* or *Aspergillus* cell, preferably of the species *Fusarium oxysporum, Fusarium venenatum, Aspergillus nidulans, Aspergillus oryzae,* or *Aspergillus niger.*

28. A recombinant vector capable of autonomous maintenance in a filamentous fungal cell, the vector comprising at least one copy of a gene encoding a dominant negative mutant of a mismatch repair protein.

29. The vector of claim 28, which comprises an AMA1 sequence.

30. The vector of claim 28, which is derived from plasmid pAAT56.

31. The vector of any of claims 28-30, wherein the dominant negative mutant is a mutant homologue of a mismatch repair protein selected from the group consisting of MutS, MutL, MutH, MSH2, MSH3, MSH6, MLH1, MLH3, PMS1, MutM, MutY, MutT, MutH, HexA, HexB, and GTBP/p160.

32. The vector of any of claims 28-31, wherein the dominant negative mutant is a mutant of the mismatch repair protein MSH2, MSH3, or MSH6.

33. The vector of any of claims 28-32, wherein the dominant negative mutant is a mutant mismatch repair protein which has a reduced ATPase activity when compared to the wildtype protein.

34. The vector of claim 33, wherein the dominant negative mutant is a mutant mismatch repair protein which has a more than 10% reduced, preferably 20%, or 30%, even 40%, more preferably 50%, or 60%, or even 70%, and most preferably 80% reduced ATPase activity when compared to the wildtype protein

35. The vector of any of claims 28-32, wherein the dominant negative mutant is a mutant mismatch repair protein which is mutated in the phosphate binding loop consensus sequence.

36. The vector of any of claims 28-32, wherein the dominant negative mutant is a mutant mismatch repair protein which is mutated in the Walker type A nucleotide binding motif.

37. The vector of any of claims 28-32, wherein the dominant negative mutant is a mutant mismatch repair protein which is mutated in the C-terminal half of the protein.

38. The vector of any of claims 28-37, wherein the gene encoding the dominant negative mutant is expressed by an inducible and/or repressible promoter.

39. The vector of any of claims 28-38, wherein the vector further comprises at least one copy of a polynucleotide encoding a polypeptide of interest; preferably an enzyme, more preferably an amylolytic enzyme, a lipolytic enzyme, a proteolytic enzyme, a cellulytic enzyme, an oxidoreductase or a plant cell-wall degrading enzyme; and most preferably an enzyme having an activity selected from the group consisting of aminopeptidase, amylase, amyloglucosidase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomeras , laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, and xylanase.

40. The vector of any of claims 28-39, which further comprises a polynucleotide encoding an error-prone polymerase, preferably a reverse transcriptase.

41. The vector of claim 40, which is replicated by the error-prone polymerase.

42. A method for preparing a filamentous fungal cell comprising a diverse polynucleotide library of interest, the process comprising:

a) introducing a polynucleotide of interest or a polynucleotide library of interest into a filamentous fungal cell as defined in any of claims 1-27 herein; and

b) cultivating the cell of (a) under conditions allowing the polynucleotide of interest or the polynucleotide library of interest to be replicated at least once.

43. The method of claim 42, wherein the filamentous fungal cell is a *Fusarium* or *Aspergillus* cell, preferably of the species *Fusarium oxysporum, Fusarium venenatum, Aspergillus nidulans, Aspergillus oryzae,* or *Aspergillus niger.*

44. The method of claim 42 or 43, wherein the polynucleotide of interest or polynucleotide library of interest comprises shuffled polynucleotides.

45. The method of any of claims 42-44, wherein the polynucleotide of interest or polynucleotide library of interest encodes a polypeptide of interest, preferably the an enzyme, and more preferably an amylolytic enzyme, a lipolytic enzyme, a proteolytic enzyme, a cellulytic enzyme, an oxidoreductase or a plant cell-wall degrading enzyme.

46. The method of any of claims 42-45, wherein the polynucleotide of interest or polynucleotide library of interest is comprised in a recombinant vector as defined in any of claims 28-41 herein.

47. The method of any of claims 42-46, wherein one or more *in vivo* recombination event(s) takes place in step b) between non-identical polynucleotides of interest.

48. Use of a filamentous fungal cell as defined in any of claims 1-27 herein for the generation of a diverse polynucleotide library of interest.

49. A process for producing a polypeptide of interest, the method comprising:

a) introducing a polynucleotide of interest or a polynucleotide library of interest into a filamentous fungal cell as defined in any of claims 1-27 herein;

b) cultivating the cell of step (a) under conditions allowing the polynucleotide of interest or the polynucleotide library of interest to be replicated at least once;

c) selecting a polynucleotide encoding a polypeptide of interest from the at least once replicated polynucleotides of step (b); and

d) expressing the selected polynucleotide in a suitable expression system to produce the polypeptide of interest.

10121.204-WO.ST25
SEQUENCE LISTING

<110> Novozymes A/S
Vind, Jesper

<120> In vivo mutation and recombination in filamentous fungi

<130> 10121.204-WO

<150> DK PA 2001 00123
<151> 2001-01-24

<150> US 60/265,300
<151> 2001-01-31

<160> 29

<170> PatentIn version 3.1

<210> 1
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 177996

<400> 1
gaatgacttg gttgacgcgt caccagtcac 30

<210> 2
<211> 25
<212> DNA
<213> Artificial sequence

<220>

<223> Primer 135640

<400> 2
cttattagta ggttggtact tcgag 25

<210> 3

<211> 37

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 135638

<400> 3
gtccccagag tagtgtcact atgtcgaggc agttaag 37

<210> 4

<211> 64

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 080399J19

<400> 4
gtatgtccct tgacaatgcg atgtatcaca tgatataatt actagcaagg gaagccgtgc 60

ttgg 64

<210> 5

<211> 59

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 051199J1

<400> 5
cctctagatc tcgagctcgg tcaccggtgg cctccgcggc cgctggatcc ccagttgtg 59

<210> 6

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 1298TAKA

<400> 6
gcaagcgcgc gcaatacatg gtgttttgat cat 33

<210> 7

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 270999J8

<400> 7
tctgtgaggc ctatggatct cagaac 26

<210> 8

<211> 27

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 270999J9

<400> 8
gatgctgcat gcacaactgc acctcag 27

<210> 9

<211> 31

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 142779

<400> 9
ttgaattgaa aatagattga tttaaaactt c 31

<210> 10

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 142780

<400> 10
ttgcatgcgt aatcatggtc atagc 25

<210> 11

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 140288

<400> 11
ttgaattcat gggtaataac tgatat 26

<210> 12

<211> 32

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 142778

<400> 12
aaatcaatct attttcaatt caattcatca tt 32

<210> 13

<211> 45

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 141223

<400> 13
ggatgctgtt gactccggaa atttaacggt ttggtcttgc atccc 45

<210> 14

<211> 44

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 141222

<400> 14
ggtattgtcc tgcagacggc aatttaacgg cttctgcgaa tcgc 44

<210> 15

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 19671

<400> 15
ctcccttctc tgaacaataa accc 24

<210> 16

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 220900j4

<400> 16
tcccatgtta ggaccagtaa tgat 24

<210> 17

<211> 66

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 991213j5

<400> 17
cctctagatc tcgagctcgg tcaccggtgg cctccgcggc cgctgcgcca ggtgtcagtc 60

accctc 66

<210> 18

<211> 39

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 220900j1

<400> 18
actggtccta acatgggagc gaaatcgact tatattcgc 39

<210> 19

<211> 39

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 220900j2

<400> 19
actggtccta acatgggagg tgcgtcgact tatattcgc 39

<210> 20

<211> 39

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 220900j3

<400> 20
actggtccta acatgggagg taaagcgact tatattcgc 39

<210> 21

<211> 51

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 16062000j2

<400> 21
ggggaccact ttgtacaaga aagctgggtc ctagatctcg agctcggtca c 51

<210> 22

<211> 50

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 1907200OJ1

<400> 22
ggggacaagt ttgtacaaaa aagcaggctt ctctgaacaa taaaccccac 50

<210> 23

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Primer U

<400> 23
gggaagctgc caggccccag tgtcagcctc ctatgctc 38

<210> 24

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Primer L-G.T

<400> 24
gagcatagga ggctgacatt ggggcctggc agcttccc 38

<210> 25

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Primer L-G.A

<400> 25
gagcatagga ggctgacaat ggggcctggc agcttccc 38

<210> 26

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Primer L-G.G

<400> 26
gagcatagga ggctgacagt ggggcctggc agcttccc 38

<210> 27

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Primer L-A.C

<400> 27
gagcatagga ggctgacacc ggggcctggc agcttccc 38

<210> 28

<211> 40

<212> DNA

<213> Artificial sequence

<220>

<223> L-T.G

<400> 28
gagcatagga ggctgacact gtggggcctg gcagcttccc 40

<210> 29

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Primer L-HOM

<400> 29
gagcatagga ggctgacacc ggggcctggc agcttccc 38

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 02/00050

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C12N 15/80 // C 12 N 5/10, C 12 N 15/01
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA, BIOSIS, MEDLINE, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6146894 A (N. NICOLAIDES ET AL.), 14 November 2000 (14.11.00), column 3, line 51 - column 6, line 26	1-2,4-7, 11-13,15,28, 31-34,38
Y	column 3, line 51 - column 6, line 26	3,14,16-27, 29-30,39-49
A	column 3, line 51 - column 6, line 26	8-9,10,35-37
	--	
Y	WO 0050567 A1 (NOVO NORDISK A/S), 31 August 2000 (31.08.00), page 2, line 19 - line 27	3,14,16-27, 29-30,39-49
	--	

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 May 2002	16-05-2002
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Frida Plym Forshell/AE Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Information on patent family members

01/05/02

International application No.
PCT/DK 02/00050

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6146894 A	14/11/00	NONE	
WO 0050567 A1	31/08/00	AU 2536400 A	14/09/00
		CN 1341146 T	20/03/02
		EP 1157095 A	28/11/01